Sequence Control Tokens: Comprehensive Comparison

SOS Token - Generation Initialization

| [SOS] | The | cat | runs | fast |

Function: Initialization

Conditions first token generation

EOS Token - Generation Termination

| Task | complete | successfully | [EOS] |

Function: Termination

Signals sequence completion
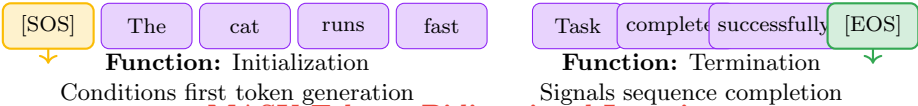
MASK Token - Bidirectional Learning

| The | [MASK] | runs | very | [MASK] |

Function: Prediction

Uses bidirectional context
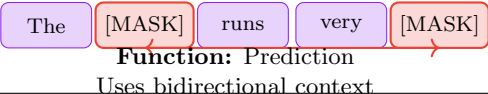
Comparative Analysis

Aspect	SOS Token	EOS Token	MASK Token
Position	Start only	End only	Any position
Context	No context	Left context	Bidirectional
Training	Autoregressive	Autoregressive	MLM
Purpose	Initialization	Termination	Prediction
In Output	Removed	Optional	Replaced
Models	GPT, T5	GPT, T5	BERT, RoBERTa

Usage Patterns in Different Architectures

GPT-style Models

Autoregressive Generation

| [SOS] | token | [EOS] |

Applications:

Text generation, Translation, Dialogue systems

BERT-style Models

Bidirectional Understanding

| token | [MASK] | token |

Applications:

Language understanding, Classification, QA